

INSTITUTO SUPERIOR MIGUEL TORGA

Escola Superior de Altos Estudos

**Contributo para o Estudo das Propriedades Psicométricas do
Teste do Desenho do Relógio numa Amostra da População
Portuguesa**

Luís Estêvão Morais Bonito Basto de Oliveira

Dissertação de Mestrado em Psicologia Clínica
Ramo de Especialização em Terapias Cognitivo-Comportamentais

Coimbra, 2016



Contributo para o Estudo das Propriedades Psicométricas do Teste do Desenho do Relógio numa Amostra da População Portuguesa

Luís Estêvão Morais Bonito Basto de Oliveira

Dissertação Apresentada ao ISMT para a Obtenção do Grau de Mestre em Psicologia Clínica
Ramo de Especialização em Terapias Cognitivo-Comportamentais
Orientadora: Professora Doutora Helena Espírito Santo, Professora Auxiliar, Instituto
Superior Miguel Torga

Coimbra, abril de 2016

Agradecimentos

À minha família nomeadamente à minha filha Constança, minha mulher Marina e aos meus Pais por todo o apoio, compreensão, motivação e suporte dado.

À Professora Doutora Helena Espirito Santo pela sua disponibilidade, saber e dedicação que sempre disponibilizou e ainda aos meus amigos e colegas que me apoiaram neste desígnio

Resumo

Objetivo: Estudar as propriedades psicométricas do Teste do Desenho do Relógio numa amostra da população portuguesa.

Método: A amostra é constituída por 749 pessoas (345 homens e 404 mulheres), com idades compreendidas entre os 18 e os 94 anos. Todos os participantes preencheram uma declaração de consentimento informado e uma bateria de testes neuropsicológicos.

Resultados: A média no Teste do Desenho do Relógio foi de 4,10 ($DP = 1,25$). Os resultados demonstraram que as variáveis sociodemográficas (idade, escolaridade, profissão, regiões e tipologia de áreas urbanas) apresentaram ter influência significativa nas pontuações do Teste do Desenho do Relógio. A confiabilidade e a estabilidade temporal revelaram-se adequadas.

Conclusão: Os dados do nosso estudo indicam que o Teste do Desenho do Relógio pode ser utilizado pela população portuguesa.

Palavras-chave: Teste do Desenho do Relógio; avaliação neuropsicológica; estudos psicométricos; adultos.

Abstract

Purpose: To study the psychometric properties of the Clock Drawing Test in a Portuguese sample.

Method: The sample consists of 749 people (345 men and 404 women), aged between 18 and 94 years. All participants filled an informed consent form and a battery of neuropsychological tests.

Results: The average in Clock Drawing Test was 4.10 ($SD = 1.25$). The results showed that the sociodemographic variables (age, education, profession, region, and typology of urban areas) showed significant influence on Clock Drawing Test scores. The reliability and temporal stability of Clock Drawing Test proved appropriate.

Conclusion: The data from our study indicate that the Clock Drawing Test may be used by the Portuguese population.

Keywords: Clock Drawing Test; neuropsychological assessment; Psychometric studies; adults.

Introdução

O Teste do Desenho do Relógio (TDR) tem sido largamente utilizado como ferramenta de avaliação neuropsicológica para rastreio cognitivo, tanto em contexto clínico, como de investigação (Bennasar, Setchi, Bayer e Hicks, 2013; Duro, Freitas, Alves, Simões e Santana, 2012; Hubbard et al., 2008; Kaplan, 1990; De Paula, Miranda, Moraes e Malloy-Diniz, 2013b; Shulman, 2000; Tranel, Rudrauf, Vianna e Damasio, 2008).

Na última década tornou-se mais frequente o seu uso, dado que realiza de forma rápida a avaliação cognitiva (Hubbard et al., 2008; Kim, Lee, Choi, Sohn e Lee, 2009; Shulman, 2000), sobretudo para a deteção precoce de défice cognitivo (Shulman, Shedletsky e Silver, 1986) ou demência (Bennasar et al., 2013; Eknoyan, Hurley e Taber, 2012; Ismail et al., 2013; Shoyama et al., 2011; Shulman, 2000). Neste contexto, alguns estudos têm relatado que o uso deste teste é igualmente útil e eficaz no diagnóstico e classificação da gravidade da encefalopatia hepática, prevendo resultados ao nível da reabilitação após a lesão cerebral traumática (Edwin, Peter, John, Eapen e Graham, 2011; Writer, Schillerstrom, Regwan e Harlan, 2010).

O TDR avalia as funções cognitivas, tais como a habilidade viso espacial, a função executiva, a compreensão e a memória semântica (Leyhe, Erb, Milian, Eschweiler, Ethofer e Grodd e Saur, 2009; Shulman, 2000). Muita da informação que existe sobre o conceito de relógio é abstrata e concetualmente complexa, por exemplo, o significado da disposição espacial do mostrador de um relógio, o significado dos ponteiros e os respetivos comprimentos. Contudo, é a natureza multifatorial dos processos por detrás do ato de desenhar um relógio que o torna bastante sensível a défices cognitivos (Tuokko e O'Connell, 2006).

Shulman (1986, 2000), 14 anos após ter constatado que existia um problema de padronização do sistema de pontuação do TDR, da formação que o avaliador devia possuir para o administrar e em que nível o TDR deveria ser realizado, procedeu a uma revisão dos estudos efetuados entre 1983 e 1998 que utilizaram o TDR. O autor refere ainda que existem vários sistemas para pontuar o TDR (em 4 classes de 10 pontos, de 20 pontos, de 6 pontos, de 5 pontos de 10 pontos mais erros de 4 + 4 quadrantes, etc.) contudo utiliza preferencialmente o de 5 pontos. No estudo mais recente Shulman (2000), destaca que o TDR é usado em complemento com o Mini Exame do Estado Mental (MMSE), entre outros, proporcionando um avanço significativo na deteção precoce da demência e assim monitorizando a mudança cognitiva. Também notou que um sistema de pontuação simples (5 pontos) com ênfase nos

aspectos qualitativos do desenho do relógio, maximiza a sua utilidade (Shulman, 2000). O mesmo refere ainda que a maioria das investigações que utilizavam o TDR apresentavam altas medidas de sensibilidade e especificidade, havendo grande aceitabilidade por parte dos pacientes e dos clínicos. Apesar da sua importância, tem sido referida a necessidade de existirem mais estudos que avaliem as características psicométricas, sobretudo no que diz respeito à validade do instrumento (Fuzikawa, Lima-Costa, Uchoa, Barreto e Shulman, 2003).

A principal limitação do TDR, em prática clínica, advém da falta de um simples e objetivo sistema de pontuação, o que dificulta a avaliação e interpretação. Infelizmente, nem as instruções, nem as avaliações são uniformes (Shulman et al., 1986; Patocskai et al., 2011). Apesar das várias vantagens do TDR, existem ainda muitos aspetos importantes que requerem um estudo mais aprofundado. No que respeita às características psicométricas do instrumento, Pinto e Peters (2009) referem que diversos estudos mostraram que o TDR apresenta uma alta confiabilidade interobservador e teste-reteste (p. ex., Hubbard et al., 2008). A correlação com outros testes de triagem foi estatisticamente significativa na maioria dos estudos, mas os resultados foram influenciados pela idade, educação e idioma (Pinto e Peters, 2009). Os autores concluíram que o TDR tem características adequadas e um bom método para a deteção de demência moderada e grave.

Tuokko, Hadjistavropoulos, Rae e O'Rourke (2000) realizaram uma pesquisa cujo objetivo foi examinar as características psicométricas de cinco métodos de pontuação do TDR: o método de Doyon pontua entre 0 e 20, em termos do desenho, difere dos restantes pelo facto de apresentar um círculo já com o número 12 e um ponto no local onde devem iniciar os ponteiros, Tuokko de 0 a 41, Watson de 0 a 7 e Wolf-Klein de 1 a 10. Os menos fidedignos foram os de Doyon e Watson. Este estudo validou todas as exigências para a adequação da tarefa do desenho do relógio para a identificação de demência, independentemente do método de pontuação aplicado. A principal descoberta deste estudo foi que os diferentes procedimentos de pontuação podem variar ao nível da confiabilidade, sensibilidade, especificidade, bem como a identificação dos grupos de diagnóstico (Tuokko et al., 2000).

Ainda sobre os métodos de pontuação do TDR, Kørner, Lauritzen, Nilsson, Lolk e Christensen (2012) realizaram um estudo que avaliou cinco métodos para pontuar o TDR, pontuados de 0 - 5; 0 - 2; 0 - 1; 1 - 10 e 1 - 6. A amostra abrangeu um total de 72 pacientes em ambulatório e 29 controlos saudáveis, sendo avaliados em três fases. Na primeira visita, utilizaram-se procedimentos de diagnóstico e avaliação através das Impressões Clínicas

Globais (CGI), bem como da Deterioração Global. Seguidamente, foram administrados aos pacientes o *Geriatric Depression Scale*, o TDR e o MMSE, sendo o TDR e o MMSE repetidos na segunda visita. Já o TDR, o CGI e o GDS foram repetidos na terceira e última visita. No final, o TDR foi submetido a avaliação com base em todos os métodos de pontuação em estudo. Para aferir o grau de adequação do TDR como ferramenta de triagem, analisaram-se os resultados. Os autores constataram que as correlações entre o TDR e as outras escalas eram estatisticamente significativas. Concluíram igualmente que a eficácia, a facilidade de administração e a velocidade foram as características mais visíveis nos instrumentos, que também se correlacionaram fortemente com a frequência de utilização. Quanto aos métodos de pontuação os resultados foram semelhantes, com valores mais elevados no método de 0 - 10 apesar de outros considerarem que o de 0 - 1 é o melhor pela rapidez de cotação (Kørner et al., 2012).

No mesmo sentido, Hubbard e colaboradores (2008) realizaram uma pesquisa utilizando três sistemas de pontuação do TDR, em 207 idosos “cognitivamente normais”. Os três sistemas utilizados que foram desenvolvidos por Mendez, Freund e Cahn correlacionaram-se significativamente, requerendo pouco tempo à sua aplicação, apresentando uma alta confiabilidade entre avaliadores. Foram encontradas diferenças estatisticamente significativas na pontuação no TDR com base na idade. Também descobriram que o desempenho no TDR “normal” inclui uma distribuição mais ampla dos resultados. Assim, estes resultados podem ser úteis para os clínicos que desejam saber se os seus pacientes realizam o TDR no mesmo intervalo de idosos cognitivamente normais (Hubbard et al., 2008).

Um estudo prospetivo realizado entre 1991 e 2009 por Bennasar e colaboradores (2013) avaliou 47 características possíveis (p. ex., sequência de números, divisão em horas, divisão em minutos, que mão utilizou, se utilizou as duas, etc.) na análise das pontuações do TDR.

De Paula e colaboradores (2013b) realizaram um estudo cujo objetivo visava avaliar a influência das funções executivas, o estado cognitivo global, o processamento viso espacial e o conhecimento semântico no TDR. O protocolo incluiu diferentes medidas para cada domínio cognitivo, todas validadas, nomeadamente o TDR, a Bateria de Avaliação Frontal, o MMSE, a Construção com Palitos e um teste de nomeação (TN-LIN), aplicando-os a 53 pacientes com défice cognitivo leve, 60 com Alzheimer e 57 idosos saudáveis. O domínio cognitivo mais relacionado com o desempenho no TDR foi o funcionamento executivo, medido pela pontuação total da FAB, seguindo-se a cognição geral, o processamento viso

espacial e o conhecimento semântico. Os outros referem que o TDR é uma medida multidimensional complexa, e por isso deve ser interpretada juntamente com outros testes cognitivos na avaliação neuropsicológica (De Paula et al., 2013b).

Paganini-Hill e Clark (2011) também testaram hipóteses relacionadas com o desempenho na tarefa do TDR, mais concretamente se diminui ao longo do tempo e se o mau desempenho está associado ao desenvolvimento de demência e sobrevivência reduzida. Os participantes foram convidados a preencher um TDR e o estado de demência foi determinado através da utilização de avaliações presenciais, questionários, dados dos hospitais e atestados de óbito. Os resultados demonstraram que em 4.842 participantes (média de idades de 80 anos) que completaram o TDR, o desempenho diminuiu ao longo do tempo. Os resultados médios foram menores em indivíduos sem demência, mas que foram identificados posteriormente como dementes, em comparação com aqueles que permaneceram sem demência. Um baixo índice no TDR foi associado a um risco acrescido de 28% de demência e 13% de sobrevivência reduzida. Concluíram assim que o TDR pode ajudar a prever o declínio cognitivo e a futura incapacidade nos idosos (Paganini-Hill e Clark, 2011).

Para que o TDR possa ser utilizado adequadamente como prova neuropsicológica ou instrumento de rastreio de défice cognitivo, é essencial conhecer o tipo de desempenho em indivíduos saudáveis. Por isso, foi realizado um estudo que analisou a influência de variáveis sociodemográficas (idade, género, nível educacional, estado civil, situação profissional, região geográfica, localização geográfica e área de residência) e de saúde (queixas subjetivas de memória do participante e avaliadas pelo cuidador, sintomatologia depressiva e história familiar de demência) no TDR, utilizando o sistema de cotação do TDR de 18 pontos, numa amostra representativa da população portuguesa ($n = 630$), cognitivamente saudável e com um mínimo de 25 anos (Duro et al., 2012). A pontuação no TDR apresentava uma relação linear com a idade e a escolaridade, explicando 26% da variância total dos resultados. A variável género revelou um efeito significativo no resultado total na prova, no entanto, apresentou um valor reduzido. Das variáveis de saúde estudadas, apenas as queixas subjetivas de memória dos participantes revelaram um efeito significativo, embora baixo, no resultado no sistema de cotação selecionado. Constatou-se que as variáveis idade e escolaridade, constituem os principais critérios para o estabelecimento de dados normativos para a população portuguesa, bem como para o desenvolvimento de pontos de corte para o défice cognitivo ligeiro e demência (Duro et al., 2012).

De Paula e colaboradores (2013a) identificaram a educação formal como um fator crítico no desempenho neuropsicológico, tendo um grande impacto no desempenho de testes

cognitivos, podendo enviesar a interpretação dos resultados. Todavia, ainda há poucos estudos que avaliaram as propriedades psicométricas, principalmente o critério de validade e testes neuropsicológicos para pacientes com baixa escolaridade. A investigação dos autores mostra uma forte evidência relativamente à validade de um protocolo neuropsicológico para a avaliação cognitiva de idosos com um baixo nível de escolaridade. As medidas são válidas para a avaliação das funções executivas, linguagem, memória e habilidades viso espaciais. Revela igualmente uma boa precisão para o diagnóstico de pacientes com doença de Alzheimer e défice cognitivo ligeiro (De Paula et al., 2013a).

Diversos estudos têm investigado as regiões do cérebro associadas com o desempenho do TDR. No entanto, os resultados têm-se revelado inconsistentes. Matsuoka e colaboradores (2011), através de ressonância magnética identificaram relações entre cada sistema de pontuação (o TDR de Shulman, o TDR de Rouleau e o CLOX) e o volume da substância cinzenta. Nos três tipos de pontuação do TDR verificaram que os resultados do mesmo foram correlacionados positivamente com o volume de matéria cinzenta no lobo parietal direito. Todavia também observaram diferenças nos diferentes tipos de pontuação. Por exemplo, o volume de matéria cinzenta dos lobos temporais posteriores bilaterais na cotação do TDR de Shulman, no lobo temporal inferior posterior direito na cotação do TDR de Rouleau e no lobo temporal direito superior posterior na cotação de CLOX (Matsuoka et al., 2011).

Relativamente ao mau desempenho no TDR, alguns autores mencionam que está relacionado com o declínio funcional no lobo temporal esquerdo (Nagahama, Okina, Suzuki, Nabatame e Matsuda, 2005), hipocampo esquerdo (Takahashi et al., 2008), bem como no lobo parietal direito e córtex cingulado posterior direito (Lee et al., 2008). Neste sentido, e de acordo com o estudo de Kim e colaboradores (2009), as funções cognitivas e executivas estão significativamente correlacionadas com o TDR (p. ex., Paula et al., 2013b). Os autores concluíram que o grau de hiperintensidade da substância branca periventricular contribuiu mais para o comprometimento do desempenho no TDR. Este estudo sugere que a disfunção executiva através da interrupção subcortical frontal e a perda de memória podem ser responsáveis pelo agravamento no desempenho do TDR (Kim et al., 2009).

Ainda no que respeita às associações entre o desempenho do TDR e a morfologia subcortical do cérebro, Seidl e equipa (2012) estudaram 54 participantes (21 pacientes com doença de Alzheimer, 23 com défice cognitivo leve e 10 controlos saudáveis), submetendo-os a uma avaliação neuropsicológica e ressonância magnética de alta resolução. O TDR foi relacionado ao volume e forma da medida da amígdala, núcleo caudado, hipocampo, núcleo accumbens, pallidum, putâmen e tálamo. Os resultados revelaram que o baixo desempenho

no TDR foi correlacionado com alterações no hipocampo, bilateralmente e à direita no globo pálido. Logo, estas associações estão relacionadas com alterações morfométricas regionalmente específicas e não com uma atrofia geral das respectivas estruturas (Seidl et al., 2012).

O TDR, segundo Tranel e equipa (2008), tem sido utilizado tradicionalmente como o teste do "lobo parietal", mas o trabalho mais empírico com o TDR tem incidido especialmente sobre as suas características, nomeadamente a sua sensibilidade e especificidade para detetar e distinguir subtipos de demência, havendo surpreendentemente poucos estudos sobre os seus correlatos neuroanatômicos. Por isso, Tranel e equipa (2008) investigaram os correlatos neuroanatômicos do TDR, usando um grupo de 133 pacientes cujas lesões abrangiam mais de duas concavidades hemisféricas e a substância branca subjacente. Verificaram que as deficiências no TDR foram mais associadas com danos no córtex parietal direito (giro supramarginal) e à esquerda no frontal-parietal inferior do córtex opercular. A análise efetuada também mostrou que os erros viso espaciais foram predominantes em pacientes com lesão no hemisfério direito, enquanto os erros ao nível de ajustamento do tempo foram predominantes em pacientes com lesão do hemisfério esquerdo (Tranel et al., 2008).

Ao concluir a revisão da literatura, verifica-se que o TDR é um instrumento neuropsicológico útil e eficaz na realização de rastreio cognitivo em adultos e idosos. O seu uso é cada vez mais frequente, dado que deteta com rapidez a presença ou ausência de défice cognitivo. O estudo demonstra que existem vários sistemas para cotar o TDR sendo que a principal limitação deste teste advém da falta de um simples e objetivo sistema de pontuação.

Objetivos

O principal objetivo deste estudo é estudar as propriedades psicométricas do TDR numa amostra de adultos da população portuguesa. Como objetivos específicos, pretendemos: 1) analisar a diferença das pontuações do TDR em função de algumas variáveis sociodemográficas; 2) determinar a confiabilidade (através do alfa de Cronbach), a estabilidade temporal (através do teste-reteste) e a concordância entre juízes (através do Capa de Cohen); 3) determinar a validade convergente, relacionando o TDR com outros instrumentos utilizados na bateria de testes.

Métodos

Participantes

Recrutámos 935 pessoas voluntárias usando o método de amostragem não-probabilística por conveniência, tendo sido selecionados por estarem disponíveis e pelo auxílio na divulgação a outras pessoas (Hill e Hill, 2000). O recrutamento teve lugar na comunidade, tendo 6 pessoas recusado participar (0,6%). As pessoas não receberam qualquer compensação financeira por participar, no entanto foi dada a opção de receber os resultados e o seu significado, 64 participantes assim o solicitaram (6,8%).

Os critérios de seleção incluíram: a) ser capaz de ler e escrever em português; b) ter nacionalidade portuguesa ou viver em Portugal há mais de 5 anos; c) ter mais de 50% de escolaridade realizada em Portugal; d) ter pontuação igual ou superior a 20 no Teste de Memória de 15-Item de Rey (Boone, Salazar, Lu, Warner-Chacon e Razani, 2002); f) ter pontuação abaixo de 40 na Escala de Autoavaliação de Ansiedade de Zung, sendo o ponto de corte para a ansiedade clinicamente significativa (Serra, Ponciano e Relvas, 1982).

Os participantes da nossa amostra foram estratificados de acordo com a idade, sendo formados oito grupos etários: 18-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-84 e 85-94 anos.

O estado civil foi estratificado segundo duas categorias: sem companheiro e com companheiro.

Os participantes da nossa amostra foram também estratificados de acordo com o nível de escolaridade. Foram formados seis grupos: sem escolaridade, 1º ciclo do ensino básico, 2º ciclo do ensino básico, 3º ciclo do ensino básico, ensino secundário e ensino superior. Entre os participantes, 1 sujeito encontrava-se sem informação.

Quanto à profissão, os participantes foram estratificados segundo duas categorias: manual (p. ex., agricultores) e intelectual (p. ex., pessoal administrativo). Dos participantes, 110 sujeitos encontravam-se sem informação.

A Nomenclatura das Unidades Territoriais para fins estatísticos (NUTS; Ministério das cidades, Ordenamento do Território e Ambiente, 2002) foi estratificada de acordo com a variável regiões, com base em três categorias: Norte; Centro; Sul e Regiões Autónomas. Entre os participantes, 1 sujeito encontrava-se sem informação.

A tipologia de áreas urbanas foi avaliada de acordo com a zona de residência. Nomeadamente se o participante vive em meio urbano, misto ou rural (Instituto Nacional Estatística, 2014). Dos participantes, 1 sujeito encontrava-se sem informação.

Instrumentos

Teste do Relógio (Shulman et al., 1986). O TDR foi adaptado com base na cotação de Shulman, Pushkar Gold, Cohen e Zuccherro (1993). O TDR encontra-se vocacionado para avaliar as funções executivas e a atenção seletiva de uma forma rápida e segura. O tempo de aplicação é de um minuto. Entrega-se uma folha A4 ao participante com um círculo desenhado com 9,5 cm de diâmetro e solicita-se que este desenhe os números no círculo para que se pareça com um relógio, a seguir solicita-se que ponha os ponteiros e que marque 11h e 10 m. A cotação varia entre zero e cinco sendo o ponto de corte para défice de < 4 . Como pontuação temos: Representação muito incorreta do relógio (zero pontos); Desorganização visoespacial grave (um ponto); Desorganização visoespacial moderada (dois pontos); Erros na indicação da hora-minutos (três pontos); Erros visoespaciais leves (quatro pontos); Relógio perfeito (cinco pontos) (Shulman et al., 1993). Houve necessidade de criar uma tipologia diferente para a pontuação com base em seis itens para a análise estatística. A tipologia foi criada tendo por base os critérios de Shua-Haim, Koppuzha, Shua-Haim e Gross (1997). Sendo, Item 1 (presença dos 12 números, independentemente da disposição especial); Item 2 (disposição especial do 3, 6, 9 e 12); Item 3 (disposição especial do 1, 2, 4 e 5); Item 4 (disposição especial do 7, 8, 10 e 11); Item 5 (presença dos ponteiros da hora e dos minutos, independentemente da disposição especial); Item 6 (horas exatas) (Shua-Haim et al., 1997). No nosso estudo obtivemos um alfa de Cronbach de 0,73.

Teste de Memória de 15-Item de Rey (15-IMT, *Rey 15-Item Memory Test*; Boone et al., 2002). O 15-IMT é um teste que pretende verificar se os participantes simulam ou não e assim sendo foi usado neste estudo para detetar a simulação e o esforço insuficiente (Reznek, 2005; Simões et al., 2010). O teste é constituído por duas tarefas: o Ensaio de Evocação Livre e o Ensaio de Reconhecimento. Na primeira tarefa é apresentado ao participante uma folha com 15 elementos distintos agrupados em cinco categorias. De seguida, o participante deve memorizar esses elementos (em dez segundos) e depois terminado esse tempo retira-se a folha e é pedido ao participante que reproduza a informação observada para uma folha branca. Na segunda tarefa é apresentada uma nova folha com 30 elementos: os 15 elementos que o participante observou anteriormente misturados com outros 15 elementos. O objetivo passa por identificar os elementos previamente observados na primeira tarefa (Griffin, Glassmire, Henderson e McCann, 1997; Reznek, 2005; Simões et al., 2010; Slick, Tan, Strauss e Hultsch, 2004). O alfa de Cronbach deste instrumento foi de 0,79, sendo que 319 pessoas responderam a este instrumento.

Escala de Autoavaliação de Ansiedade de Zung (SAS de Zung, *Zung Self-Rating Anxiety Scale*; Serra et al., 1982). A SAS de Zung pretende distinguir entre participantes que sentem ansiedade no momento da avaliação dos que não sentem (Serra et al., 1982), daí a sua presença no nosso estudo. A SAS de Zung é uma escala que discrimina bem entre a população normal e uma população doente com sujeitos considerados ansiosos. É uma escala que mede a ansiedade-estado, ou seja, a ansiedade que é sentida no momento da avaliação. A SAS de Zung é constituída por 20 itens (escala Likert) onde os itens são cotados de 1 a 4, exceto os itens 5, 9, 13, 17, 19 que são cotados em sentido inverso (4, 3, 2, 1). A pontuação varia entre 20 e 80. A SAS de Zung mede quatro aspetos da ansiedade: ansiedade cognitiva, ansiedade motora, ansiedade vegetativa e ansiedade do sistema nervoso central. Quanto mais alto for o valor obtido maior será a ansiedade-estado. Um sujeito tem uma grande probabilidade de sofrer de ansiedade clinicamente significativa quando a escala assume valores acima de 40 (Serra et al., 1982). O alfa de Cronbach deste instrumento foi de 0,76, ao qual responderam 319 pessoas.

Bateria de Avaliação Frontal (FAB, *The Frontal Assessment Battery*; Dubois, Slachevsky, Litvan e Pillon, 2000). O FAB foi escolhido pois mede o funcionamento executivo que é necessário para o estudo da validade convergente. O teste indica a presença ou não de disfunção cognitiva e a sua gravidade. O FAB é constituído por um conjunto de subtestes, onde cada subteste tem uma pontuação entre 0 e 3. Os subtestes analisam competências relacionadas com os lobos frontais: concetualização, flexibilidade mental, programação, sensibilidade à interferência, controlo inibitório e autonomia ambiental. A pontuação final varia entre 0 e 18, onde a pontuação máxima é o melhor resultado que o participante pode obter (Appollonio et al., 2005; Beato et al., 2012; Dubois et al., 2000; Lima, Meireles, Fonseca, Castro e Garrett, 2008). O alfa de Cronbach deste instrumento foi de 0,58, onde 322 pessoas responderam.

Figura Complexa de Rey (FCR, *Rey Complex Figure Test*; Rey, 2002). Este teste avalia a aptidão viso-espacial e a memória visual (Lezak, Loring e Howieson, 2004). A FCR divide-se em três momentos de avaliação: a cópia, a memória imediata (3 minutos) e a memória diferida (20 minutos). Na cópia, a FCR é mostrada ao participante de forma a que este possa reproduzir uma cópia da FCR. Na memória imediata, sem ser avisado com antecedência é solicitado ao participante a realização da imagem de memória (evocação). Na memória diferida é solicitado, de novo, ao participante que reproduza de memória o desenho da FCR. A correção e pontuação é baseada em diferentes tipos de cópia: Tipo I (construção sobre a armação), Tipo II (reprodução dos detalhes englobados na armação), Tipo III

(contorno geral), Tipo IV (justaposição de detalhes), Tipo V (fundo confuso), Tipo VI (esquema familiar) e Tipo VII (garatuja) (Lezak et al., 2004; Rey, 2002). A pontuação oscila entre 0 e 36 pontos, sendo baseados em 18 itens. A pontuação de cada item pode variar entre 2 pontos (elemento correto e bem situado), 1 ponto (elemento correto e mal situado) ou 1 ponto (elemento bem situado deformado ou incompleto mas reconhecível), 0,5 pontos (elemento mal situado deformado ou incompleto, mas reconhecível) e 0 pontos (elemento irreconhecível ou ausente) (Lezak et al., 2004; Rey, 2002). O alfa de Cronbach deste instrumento foi de 0,76, ao qual responderam 346 pessoas.

Análise Estatística

Para a análise e tratamento dos dados, utilizámos o Programa Estatístico *Statistical Package for the Social Sciences* (IBM SPSS Statistics, versão 21.0 para Macintosh Mavericks, SPSS, 2012).

As análises foram iniciadas com estatísticas descritivas para a pontuação total do TDR incluindo frequências, percentagens, médias, desvios-padrão. Realizámos o teste do qui-quadrado da aderência para explorar a proporção de casos que caem em cada categoria de cada variável e comparar essa proporção com valores hipotéticos (Pallant, 2011).

Com recurso ao teste de Shapiro-Wilk¹ e das medidas de assimetria e de curtose foi verificada a distribuição das pontuações do TDR através da análise da normalidade. Estes testes avaliam a hipótese nula de que a distribuição dos dados é normal (Pallant, 2011).

Foi utilizado o teste *t*/ANOVA e as correlações de Pearson para explorar os efeitos das características sociodemográficas (idade, sexo, estado civil, escolaridade, profissão, regiões e tipologia de áreas urbanas) sobre o desempenho do TDR. Para o teste *t* de amostras independentes, um poder de 0,95, um alfa de 0,05 e um tamanho de efeito de 0,5 (*d* de Cohen), a amostra teria de ter 210 sujeitos. Quanto à ANOVA, um poder de 0,95, um alfa de 0,05 e um tamanho de efeito de 0,25 (*f* de Cohen), a amostra teria de ter entre 343 e 360 sujeitos (Faul, Erdfelder, Lang e Buchner, 2007a, 2007b).

Através do teste da análise da variância (ANOVA) determinámos a homogeneidade das variâncias segundo o teste de Levene. No caso da existência de homogeneidade ($p >$

¹ Aquando da análise da normalidade, o teste de Shapiro-Wilk tem sido reportado como sendo mais potente que o teste de Kolmogorov-Smirnov (Razali e Wah, 2011). Existem problemas quanto à probabilidade da rejeição da hipótese nula devido à possibilidade de esta aumentar na presença de uma amostra grande (superior a 300 participantes). Assim sendo, os testes de normalidade podem não ser confiáveis e deve-se usar as medidas de assimetria e de curtose (Kim, 2013).

0,05) recorreremos ao teste *post hoc* Hochberg, não existindo será utilizado o teste *post hoc* Games-Howell, ambos com a correção de Bonferroni (p / n° de comparações par-a-par).

Relativamente à análise das propriedades psicométricas, determinámos a consistência interna através do alfa de Cronbach, usando uma tipologia criada por nós tendo por base os critérios de Shua-Haim et al. (1997). Na análise de consistência interna para um alfa de 0,05, um poder de 0,95 e um alfa de Cronbach esperado mínimo de 0,7, o tamanho da amostra necessária seria de 24 sujeitos (Bonett, 2002; Chang, 2014).

Utilizámos a análise correlacional para a validade convergente com alguns dos testes utilizados na bateria de testes. Quanto à análise correlacional, um poder de 0,95, um alfa de 0,05 e um tamanho de efeito de 0,5 (q de Fisher), a amostra teria de ter 214 sujeitos (Faul et al., 2007a, 2007b).

Para a análise da estabilidade temporal teste-reteste utilizámos correlações de Pearson e o teste t de Student para amostras emparelhadas.

Calculámos o coeficiente Capa de Cohen uma medida de concordância interavaliador que mede o grau de concordância entre juízes. Os valores variam entre 0 e 1, onde 1 representa uma concordância perfeita. Um valor de 0,5 representa uma concordância moderada, acima de 0,7 uma boa concordância e acima de 0,8 representa uma muito boa concordância (Peat, 2002).

Para a magnitude do efeito através do d de Cohen (Espírito-Santo e Daniel, 2015) podemos pontuar este como: pequeno ($< 0,2 - 0,49$); médio ($0,5 - 0,79$); grande ($0,8 - 1,29$); muito grande ($> 1,30$).

Procedimentos

Todos os participantes preencheram uma declaração de consentimento informado (Apêndice A) de acordo com a Declaração de Helsínquia garantindo a confidencialidade e o anonimato dos dados. Como parte do protocolo, os participantes tiveram de completar uma bateria de testes: o TDR (Shulman et al., 1986), o 15-IMT (Boone et al., 2002), a SAS de Zung (Serra et al., 1982), o Inventário de Lateralidade de Edinburgh (EHI, Edinburgh Handedness Inventory; Oldfield, 1971), a FAB (Dubois et al., 2000), o Teste Stroop (Stroop, 1935), a Forma Geral das Matrizes Progressivas de Raven (FG-MPR; *Raven's Standard Progressive Matrices*, Raven, Raven e Court, 2000) e a FCR (Rey, 2002). Os testes foram administrados com duração aproximada de uma hora em espaços reservados sem elementos distratores, entre agosto de 2015 e outubro de 2015.

Resultados

Da nossa amostra inicial de 935 participantes foram excluídos 151 participantes com idades compreendidas entre os 7 e os 17 anos e outros 12 participantes devido à sua nacionalidade (uma de nacionalidade suíça, dez de nacionalidade brasileira e uma de nacionalidade angolana), visto o nosso estudo se destinar meramente à população portuguesa.

Foram eliminados vinte e três participantes “simuladores” (pontuação no teste 15-IMT inferior a 19) uma vez que as suas pontuações no TDR foram significativamente [$t(23,625) = 3,86; p < 0,05$] inferiores ($M = 3,43; DP = 1,27$) aos participantes que não simulavam ($M = 4,48; DP = 0,90$). A magnitude da diferença foi grande (d de Cohen = 1,13).

Quanto à ansiedade, os resultados demonstraram que a mesma não afetava o desempenho dos participantes na execução do teste do TDR [$t(81,379) = 1,21; p > 0,05$]. A magnitude da diferença foi pequena (d de Cohen = 0,20), pelo que optámos por manter estes sujeitos.

A amostra ficou assim reduzida a um total de 749 participantes, 345 eram do sexo masculino (46,1%) e 404 do sexo feminino (53,9%). As idades variaram entre 18 e 94 anos ($M = 40,31; DP = 19,49$).

Relativamente ao estado civil, 298 dos participantes eram casados ou viviam em união de facto (39,8%) enquanto 291 eram solteiros (38,9%).

O nível de escolaridade dos participantes variou entre a iliteracia (2,4%) e os 28 anos de escolaridade ($M = 12,77; DP = 5,05$).

Em relação à profissão a amostra destacou-se mais na categoria Estudante (33,6%), no Pessoal de serviços e vendedores (15,8%) e em Técnicos de nível intermédio (13,5%).

Relativamente às NUTS, 636 sujeitos (85,0%) viviam no Centro, 54 (7,2%) no Norte, 27 (3,6%) em Lisboa, 27 (3,6%) no Alentejo, 1 (0,1%) no Algarve e 3 (0,4%) nas regiões autónomas da Madeira e Açores.

Quanto à tipologia das áreas urbanas, 458 sujeitos (61,2%) viviam em meio urbano, 254 sujeitos (34,0%) em meio rural e 36 sujeitos (4,8%) numa zona mista, respetivamente.

Através do qui-quadrado da aderência foi possível verificar que a amostra não estava equilibrada na maioria das variáveis sociodemográficas.

Tabela 1*Caracterização Sociodemográfica (N = 749)*

		<i>n</i>	%	χ^2 ^a
Idade (<i>M</i> = 40,31; <i>DP</i> = 19,49)	18 - 24	223	29,8	345,16***
	25 - 34	148	19,8	
	35 - 44	101	13,5	
	45 - 54	110	14,7	
	55 - 64	69	9,2	
	65 - 74	28	3,7	
	75 - 84	51	6,8	
	85 - 94	19	2,5	
Sexo	Masculino	345	46,1	4,65*
	Feminino	404	53,9	
Estado Civil	Sem companheiro	377	50,3	0,03 ^{NS}
	Com companheiro	372	49,7	
Escolaridade (<i>M</i> = 12,77; <i>DP</i> = 5,05)	Sem escolaridade	18	2,4	760,93***
	1º CEB	78	10,4	
	2º CEB	46	6,1	
	3º CEB	67	9,0	
	Ensino secundário	230	30,7	
	Curso superior	310	41,3	
Profissão	Manual	88	13,8	335,48***
	Intelectual	551	86,2	
Regiões	Norte	54	7,2	1030,53***
	Centro	663	88,6	
	Sul e R. A.	31	4,1	
Tipologia	Urbano	458	61,2	357,25***
	Misto	36	4,8	
	Rural	254	34,0	

Notas: *M* = Média; *DP* = Desvio Padrão; ^a Qui-Quadrado da aderência; CEB = Ciclo do ensino básico; * $p < 0,05$; *** $p < 0,001$; ^{NS} Não Significativo; Sul e R. A. = Sul e Regiões Autônomas.

Descritivas

A média no TDR foi de 4,10 (*DP* = 1,25). A distribuição das pontuações do TDR foi simétrica (*Si* = -1,18) e mesocúrtica (*Ku* = 0,29). No entanto, o teste de Kolmogorov-Smirnov com a correção de Lilliefors e o teste de Shapiro-Wilk revelaram a violação da assunção da normalidade, tendo-se obtido um $p < 0,001$. Ainda assim e de acordo com Kim (2013) os valores da curtose (inferior a 7) e da assimetria (superior a -2) são indicativos da normalidade para amostras superiores a 300 participantes. Por essa razão, procedemos às análises

estatísticas paramétricas.

Influência das Variáveis Sociodemográficas

Na Tabela 2 é possível verificar as diferenças nas pontuações do TDR entre os grupos definidos pelas variáveis sociodemográficas, através do teste t de Student de amostras independentes ou da ANOVA.

As pontuações totais no TDR diferiram significativamente entre as oito faixas etárias [$F(7, 210) = 17,21; p < 0,001$], com um efeito médio ($\eta^2 = 0,18$) (Tabela 2).

No que diz respeito ao sexo, verificámos que não houve diferenças estatisticamente significativas [$t(744,417) = 1,27; p = 0,05$] entre homens e mulheres. A magnitude da diferença foi insignificante (d de Cohen = 0,10).

Quanto à variável estado civil, verificámos que não houve diferenças estatisticamente significativas [$t(747) = 0,25; p = 0,05$] entre participantes com e sem companheiro. A magnitude da diferença foi insignificante (d de Cohen = 0,02) (Tabela 2).

As pontuações totais no TDR diferiram significativamente entre os níveis de escolaridade [$F(5, 180) = 23,87; p < 0,001$], com um efeito médio ($\eta^2 = 0,17$).

Relativamente à análise das respostas no TDR (Tabela 2), participantes com uma profissão intelectual tiveram uma pontuação mais alta do que participantes com uma profissão manual de forma estatisticamente significativa [$t(104,769) = 2,91; p < 0,01$]. A magnitude da diferença foi pequena (g de Hedges = 0,40).

As pontuações totais no TDR diferiram significativamente entre as regiões [$F(2, 102) = 8,26; p < 0,001$], com um efeito pequeno ($\eta^2 = 0,01$).

As pontuações totais no TDR diferiram significativamente entre a tipologia de áreas urbanas [$F(2, 132) = 3,79; p < 0,01$], com um efeito pequeno ($\eta^2 = 0,01$) (Tabela 2).

Tabela 2

Diferenças nas Pontuações do Teste do Desenho do Relógio entre os Grupos Definidos pelas Variáveis Sociodemográficas (N = 749)

		<i>n</i>	<i>M</i> ± <i>DP</i>	IC 95% LI - LS	Amplitude Min - Máx
Idade <i>F</i> (7, 210) = 17,21 <i>p</i> < 0,001; $\eta^2 = 0,18$	18 - 24	223	4,44 ± 0,91	4,32 – 4,56	2 – 5
	25 - 34	148	4,39 ± 1,03	4,22 – 4,56	1 – 5
	35 - 44	101	4,26 ± 1,05	4,05 – 4,46	1 – 5
	45 - 54	110	4,08 ± 1,17	3,86 – 4,30	1 – 5
	55 - 64	69	3,86 ± 1,33	3,54 – 4,17	0 – 5
	65 - 74	28	3,43 ± 1,60	2,81 – 4,05	1 – 5
	75 - 84	51	2,90 ± 1,54	2,47 – 3,33	0 – 5
	85 - 94	19	2,05 ± 1,58	1,29 – 2,81	0 – 5
Sexo <i>t</i> (744,417) = 1,27 <i>p</i> > 0,05; <i>d</i> = 0,10	Masculino	345	4,16 ± 1,18	4,03 – 4,28	0 – 5
	Feminino	404	4,04 ± 1,30	3,92 – 4,17	0 – 5
Estado civil <i>t</i> (747) = 0,25 <i>p</i> = 0,05; <i>d</i> = 0,02	Sem companheiro	377	4,11 ± 1,26	3,98 – 4,24	0 – 5
	Com companheiro	372	4,09 ± 1,23	3,96 – 4,21	0 – 5
Escolaridade <i>F</i> (5, 180) = 23,87 <i>p</i> < 0,001; $\eta^2 = 0,17$	Sem escolaridade	18	2,00 ± 1,46	1,28 – 2,72	0 – 5
	1º CEB	78	3,14 ± 1,57	2,79 – 3,49	0 – 5
	2º CEB	46	3,59 ± 1,26	3,21 – 3,96	0 – 5
	3º CEB	67	4,09 ± 1,14	3,81 – 4,37	1 – 5
	Ensino secundário	230	4,26 ± 1,15	4,11 – 4,41	0 – 5
	Curso superior	310	4,42 ± 0,95	4,31 – 4,52	1 – 5
Profissão <i>t</i> (104,769) = 2,91 <i>p</i> < 0,01; <i>g</i> = 0,40	Manual	88	3,84 ± 1,37	3,55 – 4,13	0 – 5
	Intelectual	551	4,29 ± 1,08	4,20 – 4,38	1 – 5
Regiões <i>F</i> (2, 102) = 8,26 <i>p</i> < 0,001; $\eta^2 = 0,01$	Norte	54	4,46 ± 0,97	4,20 – 4,73	0 – 5
	Centro	663	4,05 ± 1,27	3,96 – 4,15	0 – 5
	Sul e R. A.	31	4,52 ± 0,81	4,22 – 4,81	3 – 5
Tipologia <i>F</i> (2, 132) = 3,79 <i>p</i> < 0,01; $\eta^2 = 0,01$	Urbano	458	4,21 ± 1,16	4,10 – 4,31	0 – 5
	Misto	36	4,06 ± 1,33	3,61 – 4,51	0 – 5
	Rural	254	3,93 ± 1,35	3,76 – 4,09	0 – 5

Notas: *M* = Média; *DP* = Desvio Padrão; IC 95% = Intervalo de Confiança a 95%; LI = Limite inferior; LS = Limite superior; Min = mínimo; Máx = máximo; *F* = ANOVA; *t* = teste *t* de Student; *p* = nível de significância estatística; *d* = *d* de Cohen; η^2 = eta quadrado (soma dos quadrados entre grupos / soma total dos quadrados); CEB = Ciclo do ensino básico; Sul e R. A. = Sul e Regiões Autônomas.

Comparações *a posteriori* entre pares de médias (Tabela 3), realizadas recorrendo ao teste de *post hoc* Games-Howell com a correção de Bonferroni, revelaram que as cinco faixas etárias entre os 18 e os 64 anos tiveram pontuações significativamente mais elevadas do que as idades nas faixas etárias entre os 75 e os 94 anos. Além do mais, a faixa etária dos 55 aos 64 anos também diferiram da faixa etária dos 85 aos 94 anos.

Tabela 3

Comparações Post Hoc das Pontuações do Teste do Desenho do Relógio em que houve Diferenças com a Variável Idade [$F(7, 210) = 17,21$; $p < 0,001$] ($N = 749$)

Comparação por faixa etária		Diferença média	p	d	Interpretação d
18 - 24 ($M \pm DP = 4,44 \pm 0,91$)	75 - 84	1,542	< 0,001	1,46	Efeito muito grande
	85 - 94	2,391		2,45	Efeito muito grande
25 - 34 ($M \pm DP = 4,39 \pm 1,03$)	75 - 84	1,490	< 0,001	1,26	Efeito grande
	85 - 94	2,339		2,11	Efeito muito grande
35 - 44 ($M \pm DP = 4,26 \pm 1,05$)	75 - 84	1,355	< 0,001	1,10	Efeito grande
	85 - 94	2,205		1,93	Efeito muito grande
45 - 54 ($M \pm DP = 4,08 \pm 1,17$)	75 - 84	1,180	< 0,001	0,91	Efeito grande
	85 - 94	2,029	< 0,01	1,64	Efeito muito grande
55 - 64 ($M \pm DP = 3,86 \pm 1,33$)	85 - 94	1,802	< 0,01	1,31	Efeito muito grande

Notas: M = Média; DP = Desvio Padrão; p = nível de significância estatística; d = d de Cohen.

Comparações *a posteriori* entre pares de médias (Tabela 4), realizadas recorrendo ao teste de *post hoc* Games-Howell com a correção de Bonferroni, revelaram que as pontuações dos participantes sem escolaridade diferiram significativamente dos participantes com todos os restantes níveis de escolaridade com exceção do 1º ciclo do ensino básico. Do mesmo modo, as pontuações dos participantes do 1º ciclo do ensino básico diferiram significativamente dos participantes que possuíam o 3º ciclo do ensino básico, o ensino secundário e um curso superior. Os participantes com o 2º ciclo do ensino básico também diferiram da pontuação dos participantes com um curso superior.

Tabela 4

Comparações Post Hoc das Pontuações do Teste do Desenho do Relógio em que houve Diferenças com a Variável Escolaridade [$F(5, 180) = 23,87$; $p < 0,001$] ($N = 749$)

Comparação por nível de escolaridade		Diferença média	p	d	Interpretação d
2º CEB ($M \pm DP = 3,59 \pm 1,26$)	Sem escolaridade	1,587	$< 0,01$	1,21	Efeito grande
3º CEB ($M \pm DP = 4,09 \pm 1,14$)	Sem escolaridade	2,090	$< 0,001$	1,73	Efeito muito grande
	1º CEB	0,949		0,69	Efeito médio
Ensino secundário ($M \pm DP = 4,26 \pm 1,15$)	Sem escolaridade	2,257	$< 0,001$	1,93	Efeito muito grande
	1º CEB	1,115		0,88	Efeito grande
Curso superior ($M \pm DP = 4,42 \pm 0,95$)	Sem escolaridade	2,417	$< 0,001$	2,46	Efeito muito grande
	1º CEB	1,276		1,16	Efeito grande
	2º CEB	0,831	$< 0,01$	0,83	Efeito grande

Notas: M = Média; DP = Desvio Padrão; p = nível de significância estatística; d = d de Cohen.

Comparações *a posteriori* entre pares de médias (Tabela 5), realizadas recorrendo ao teste de *post hoc* Games-Howell com a correção de Bonferroni, revelaram que havia diferenças estatisticamente significativas nas pontuações dos sujeitos que residiram no Centro em comparação com os sujeitos que residiram no Norte e Sul.

Tabela 5

Comparações Post Hoc das Pontuações do Teste do Desenho do Relógio em que houve Diferenças com a Variável Regiões [$F(2, 102) = 8,26$; $p < 0,001$] ($N = 749$)

Comparação por regiões		Diferença média	p	d	Interpretação d
Norte ($M \pm DP = 4,46 \pm 0,97$)	Centro	0,409	$< 0,05$	0,33	Efeito pequeno
Sul ($M \pm DP = 4,52 \pm 0,81$)	Centro	0,462	$< 0,05$	0,38	Efeito pequeno

Notas: M = Média; DP = Desvio Padrão; p = nível de significância estatística; d = d de Cohen.

Comparações *a posteriori* entre pares de médias (Tabela 6), realizadas recorrendo ao teste de *post hoc* Games-Howell com a correção de Bonferroni, revelaram que as pontuações dos participantes que residiram numa zona urbana eram significativamente mais elevadas do que os participantes que residiram numa zona rural.

Tabela 6

Comparações Post Hoc das Pontuações do Teste do Desenho do Relógio em que houve Diferenças com a Variável Tipologia de Áreas Urbanas [$F(2, 132) = 3,79$; $p < 0,01$] ($N = 749$)

Comparação por tipologia		Diferença <i>média</i>	<i>p</i>	<i>d</i>	Interpretação <i>d</i>
Urbano (<i>M</i> ± <i>DP</i> = 4,21 ± 1,16)	Rural	0,280	< 0,05	0,23	Efeito pequeno

Notas: M = Média; DP = Desvio Padrão; p = nível de significância estatística; d = d de Cohen.

Propriedades Psicométricas

Confiabilidade

No presente estudo, através de uma tipologia criada por nós tendo por base os critérios de Shua-Haim et al. (1997) foi obtido um alfa de Cronbach de 0,73. O TDR apresenta uma consistência interna baixa (Murphy e Davidshofer, 2004).

Validade convergente

A Tabela 7 mostra que as correlações do TDR com a cópia, a memória imediata, a memória diferida da FCR e com a FAB foram todas significativamente positivas e baixas.

Tabela 7

Correlações entre o Teste do Desenho do Relógio, a Figura Complexa de Rey – Cópia, a Figura Complexa de Rey – Memória Imediata, a Figura Complexa de Rey – Memória Diferida e a Bateria de Avaliação Frontal ($N = 749$)

Instrumentos	Correlações				
	1	2	3	4	5
1. TDR	—	0,20**	0,23**	0,15**	0,20**
2. FCR Cópia		—	0,43**	0,25**	0,10
3. FCR Memória Imediata			—	0,51**	0,16**
4. FCR Memória Diferida				—	0,09
5. FAB					—

Notas: TDR = Teste do Desenho do Relógio; FCR = Figura Complexa de Rey; FAB = Bateria de Avaliação Frontal;

** correlação significativa $p < 0,01$.

Validade teste-reteste

Foi novamente administrada a bateria de testes a um grupo de 39 sujeitos da amostra inicial (5,2%) com o objetivo de avaliar a estabilidade temporal do TDR. O intervalo de tempo entre os dois momentos foi, em média, de 5 meses ($DP = 1,17$). Através do teste t de Student para amostras emparelhadas verificámos que não houve diferenças entre os dois

momentos [$t(38) = 1,50$; $p > 0,05$]. A diferença da magnitude foi pequena (d de Cohen = 0,24). A relação entre os dois momentos foi averiguado através do coeficiente de correlação de Pearson onde este revelou um r de 0,17 ($p > 0,05$) correspondendo a uma correlação baixa e positiva.

Coeficiente Capa de Cohen

Obtivemos um valor de 0,65 com um nível de significância de $p < 0,001$, o que significa que há uma concordância moderada elevada entre juízes (Peat, 2002).

Discussão/Conclusão

O nosso principal objetivo foi o de estudar as propriedades psicométricas, de forma a que seja possível a validação do TDR para a população Portuguesa.

Este estudo realizou-se com base numa amostra de 749 participantes de ambos os sexos, com idades compreendidas entre os 18 anos e os 94 anos, subdivididos em grupos etários de 10 anos. Ainda como características destacáveis da nossa amostra podemos referir que a maioria dos inquiridos possuem o ensino secundário ou curso superior, profissões intelectuais, são da zona centro e residem em meio urbano.

Os instrumentos aplicados foram uma adaptação da versão de Shulman do TDR, o 15-IMT, o SAS de Zung, a FAB e a FCR, todos apresentam uma boa consistência interna com exceção da FAB, que apresentou valores inferiores a 0,6. De acordo com o objetivo do estudo procederam-se a várias comparações que permitissem conhecer as propriedades psicométricas do TDR e concluímos que a idade e o nível de escolaridade influenciam de forma significativa os resultados do TDR, sendo maior nos grupos etários mais velhos. Com efeito a variação das pontuações obtidas no TDR apresenta deficits superiores nos mais velhos e menos escolarizados tal como no estudo de Pinto e Peters (2009). Neste último caso (nível de escolaridade) as maiores dificuldades situam-se em quem não tem escolaridade e quem a tem independentemente do nível de escolaridade em que se encontra.

Nas restantes variáveis sociodemográficas, sexo, estado civil, região e tipologia de áreas urbanas o efeito do tamanho das diferenças foi sempre pequeno apesar de estatisticamente significativo.

No que diz respeito às propriedades psicométricas avaliadas pela confiabilidade, validade convergente, teste-reteste e grau de concordância entre juízes, os resultados demonstram que o teste tem consistência interna e temporal e concordância entre juízes. O

estudo de Fuzikawa e colaboradores (2003) demonstra uma boa a excelente confiabilidade numa população idosa com baixíssimo nível de escolaridade, já no nosso estudo considerando que temos uma amostra mais abrangente e com um nível de escolaridade alto podemos considerar os nossos valores como positivos.

Quanto à validade convergente com os testes selecionados apesar de apresentarem correlação significativa com as memórias imediata e diferida com a cópia (FCR) e funcionamento executivo (FAB), as correlações foram baixas ou moderadas o que nos remete para utilização de outros testes na avaliação da convergência ou para o facto do TDR não avaliar estes critérios.

A principal limitação do presente estudo deve-se ao facto da amostra não ser representativa da população portuguesa, quer pela percentagem entre as camadas mais jovens da nossa amostra (29,8%) ou pelo elevado número de participantes com o ensino secundário (30,7%) e o ensino superior (41,3%) que vai ao encontro da enorme disparidade entre a profissão manual e intelectual (86,2%). Na variável regiões também verificamos uma percentagem elevada do Centro em relação às restantes, mais de 80% da amostra recai nesta categoria. Esta limitação resulta do tipo de amostragem que foi utilizado (não-probabilístico por conveniência em bola de neve). Em estudos futuros sugerimos a aplicação de um método de amostragem aleatório, pois só assim é possível universalizar os resultados obtidos a partir da amostra (Hill e Hill, 2000).

Durante a realização da bateria de testes os participantes não se encontravam a tomar medicação ou apresentavam sintomas de doença que potencialmente afetasse a realização dos testes. No entanto, é possível que houvesse participantes que sofressem de condições mínimas não diagnosticadas ou de défice cognitivo ligeiro. Em estudos futuros deve existir uma maior precisão diagnóstica dos participantes no que diz respeito à relação entre doença/toma de medicação e as pontuações totais nos testes.

Outra limitação refere-se à educação, onde esta foi operacionalizada como o número de anos do ensino regular formal concluída com sucesso. Esta abordagem é vulnerável às inúmeras modificações no sistema de educação em Portugal. Desde 2005, uma grande porção da população adulta portuguesa com baixa escolaridade matriculou-se nas “Novas Oportunidades”, para melhorar as suas qualificações e adquirir certificados de educação do ensino primário (9º ano) ou secundário (12º ano). Todavia, a equivalência atual entre escola regular e o programa de educação “Novas Oportunidades” ainda é desconhecida. Assim sendo, apenas a escolaridade regular foi creditada ao contexto de educação dos participantes.

Uma das limitações do nosso estudo nas propriedades psicométricas foi o número

relativamente pequeno de sujeitos ($n = 39$) para determinar a estabilidade temporal. Para o teste t de amostras emparelhadas, um poder de 0,95 e um alfa de 0,05, a amostra teria de ter 54 sujeitos (Faul et al., 2007a, 2007b).

A grande limitação da utilização quotidiana do TDR na prática clínica decorre da falta de um simples e objetivo sistema de pontuação e isso dificulta a sua interpretação e a sua avaliação exata (Patocskai et al., 2011). Isto leva a que os clínicos se dividam aquando de qual dos TDR usar visto a existência de várias versões com distintos sistemas de pontuações. O TDR é um teste útil para os clínicos, embora os sistemas de pontuação usados precisem de ser tidos em consideração de forma a serem revistos (Pinto e Peters, 2009), podendo assim criar um único sistema de pontuação elegível. De forma a reforçar a importância do estudo e do uso do TDR, clínicos canadianos consideram o TDR como um dos instrumentos de triagem cognitiva mais utilizados, bem como os seus colegas internacionais (Ismail et al., 2013). É um teste rápido, fácil de ser administrado e com grande aceitabilidade por parte dos participantes (Shulman et al., 2000).

Futuros estudos devem ter em consideração os vários sistemas de pontuação visto diferirem em termos de confiabilidade, sensibilidade e na identificação de grupos de diagnósticos (Tuokko et al., 2000). Diversos tipos de versões podem não estar ao dispor dos clínicos bem como a sua forma de utilização e cotação, daí a necessidade de reafirmar de novo a criação de um único sistema de pontuação, para que o TDR possa ser usado de forma eficaz.

Os dados do nosso estudo indicam que o Teste do Desenho do Relógio pode ser utilizado pela população portuguesa.

Referências Bibliográficas

- Appollonio, I., Leone, M., Isella, V., Piamarta, F., Consoli, T., Villa, M. L., ... Nichelli, P. (2005). The Frontal Assessment Battery (FAB): normative values in an Italian population sample. *Neurological Sciences*, 26(2), 108–116. <http://doi.org/10.1007/s10072-005-0443-4>
- Beato, R., Amaral-Carvalho, V., Guimarães, H. C., Tumas, V., Souza, C. P., Oliveira, G. N. e Caramelli, P. (2012). Frontal assessment battery in a Brazilian sample of healthy controls: normative data. *Arquivos de Neuro-Psiquiatria*, 70(4), 278–280. <http://doi.org/10.1590/S0004-282X2012005000009>

- Bennasar, M., Setchi, R., Bayer, A. e Hicks, Y. (2013). Feature selection based on information theory in the Clock Drawing Test. *Procedia Computer Science*, 22, 902-911. doi: 10.1016/j.procs.2013.09.173
- Bonett, D. G (2002). Sample size requirements for testing and estimating coefficient alpha. *Journal of Educational and Behavioral Statistics*, 27(4), 335–340. doi:10.3102/10769986027004335
- Boone, K. B., Salazar, X., Lu, P., Warner-Chacon, K. e Razani, J. (2002). The Rey 15-Item recognition trial: A technique to enhance sensitivity of the Rey 15-Item memorization test. *Journal of Clinical and Experimental Neuropsychology (Neuropsychology, Development and Cognition: Section a)*, 24(5), 561–573. doi: 10.1076/jcen.24.5.561.1004
- Chang, A. (2014). *StatsToDo* [Calculadora online]. Brisbane, Queensland, Austrália. Acedido em 16, fevereiro, 2016, em https://www.statstodo.com/SSiz1Alpha_Pgm.php#
- De Paula, J. J., Bertola, L., Ávila, R. T., Moreira, L., Coutinho, G., Moraes ... Malloy-Diniz, L. F. (2013a). Clinical applicability and cutoff values for an unstructured neuropsychological assessment protocol for older adults with low formal education. *PLoS ONE*, 8(9), 1-9. doi:10.1371/journal.pone.0073167.
- De Paula, J. J., Miranda, D. M., Moraes, E. N. e Malloy-Diniz., L. F. (2013b). Mapping the clockworks: what does the Clock Drawing Test assess in normal and pathological aging? *Arquivos de Neuro-Psiquiatria*, 71(10), 763-768. doi: 10.1590/0004-282X20130118
- Dubois, B., Slachevsky, A., Litvan, I. e Pillon, B. (2000). A frontal assessment battery at bedside. *Neurology*, 55(11), 1621–1626. doi: 10.1212/WNL.55.11.1621
- Duro, D., Freitas, S., Alves, L., Simões, M. R. e Santana, I. (2012). O Teste do Desenho do Relógio: Influência das variáveis sociodemográficas e de saúde na população portuguesa. *Sinapse*, 12(1), 5-12.
- Edwin, N., Peter, J. V., John, G., Eapen C. E. e Graham, P. L. (2011). Relationship between clock and star drawing and the degree of hepatic encephalopathy. *Postgraduate Medical Journal*, 87 (1031), 605-611. doi: 10.1136/pgmj.2010.108985
- Espírito-Santo, H. e Daniel, F. (2015). Calcular e apresentar tamanhos do efeito em trabalhos científicos (1): As limitações do $p < 0,05$ na análise de diferenças de médias de dois grupos. *Revista Portuguesa de Investigação Comportamental e Social*, 1(1), 3–16.
- Eknoyan, D, Hurley, R. A. e Taber, K. H. (2012). The clock drawing task: Common errors and functional neuroanatomy. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 24(3), 260-265. doi: 10.1176/appi.neuropsych.12070180

- Faul, F., Erdfelder, E., Lang, A.-G. e Buchner, A. (2007a). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. doi: 10.3758/BF03193146
- Faul, F., Erdfelder, E., Lang, A.-G. e Buchner, A. (2007b). *G*Power 3* (Versão 3.1.9.2 for Mac OS X) [Programa Informático]. Düsseldorf: Heinrich-Heine-Universität Dusseldorf. Acedido em 7, fevereiro, 2016, em <http://www.gpower.hhu.de/en.html>
- Fuzikawa, C., Lima-Costa, M. F., Uchoa, E., Barreto, S. M. e Shulman, K. (2003). A population based study on the intra and inter-rater reliability of the clock drawing test in Brazil: the Bambuí health and ageing study. *International Journal of Geriatric Psychiatry*, 18(5), 450-456. doi: 10.1002/gps.863
- Griffin, G. A. E., Glassmire, D. M., Henderson, E. A. e McCann, C. (1997). Rey II: Redesigning the Rey screening test of malingering. *Journal of Clinical Psychology*, 53(7), 757–766. doi: 10.1002/(SICI)1097-4679(199711)53:7<757::AID-JCLP14>3.0.CO;2-K
- Hill, M. e Hill, A. (2000). *Investigação por questionário*. Lisboa: Edições Sílabo.
- Hubbard, E. J., Santini, V., Blankevoort, C. G., Volkers, K. M., Barrup, M. S., Byerly, L., ... Stern, R. A. (2008). Clock drawing performance in cognitively normal elderly. *Archives of Clinical Neuropsychology*, 23(3), 295-327. doi: 10.1016/j.acn.2007.12.003
- Instituto Nacional de Estatística (2014). *Estatísticas Demográficas 2011*. Lisboa: INE.
- Ismail, Z., Mulsant, B. H., Herrmann, N., Rapoport, M., Nilsson, M. e Shulman, K. (2013). Canadian academy of geriatric psychiatry survey of brief cognitive screening instruments. *Canadian Geriatrics Journal*, 16(2), 54-60. doi: doi.org/10.5770/cgj.16.81
- Kaplan, E. (1990). The process approach to neuropsychological assessment of psychiatric patients. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 2(1), 72-87. doi: 10.1176/jnp.2.1.72
- Kim, H.-Y. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry e Endodontics*, 38(1), 52–54. doi:10.5395/rde.2013.38.1.52
- Kim, Y.-S., Lee, K.-M., Choi, B. H., Sohn, E.-H. e Lee, A. Y. (2009). Relation between the Clock Drawing Test (CDT) and structural changes of brain in dementia. *Archives of Gerontology and Geriatrics*, 48(2), 218-221. doi: 10.1016/j.archger.2008.01.010
- Kørner, E. A., Lauritzen, L., Nilsson, F. M., Lolk, A. e Christensen, P. (2012). Simple scoring of the Clock-Drawing Test for dementia screening. *Danish Medical Journal*, 59(1), 1-5.

- Lee, D. Y., Seo, E. H., Choo, I. H., Kim, S. G., Lee, J. S., Lee, D. S., ... Woo, J. I. (2008). Neural correlates of the clock drawing test performance in Alzheimer's disease: a FDG-PET study. *Dementia and Geriatric Cognitive Disorders*, 26(4), 306-313. doi: 10.1159/000161055
- Leyhe, T., Erb, M., Milian, M., Eschweiler, G.W., Ethofer, T., Grodd, W. e Saur, R. (2009). Changes in cortical activation during retrieval of clock time representations in patients with mild cognitive impairment and early Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, 27(2), 117-132. doi: 10.1159/000197930
- Lezak, M. D., Loring, D. W. e Howieson, D. B. (2004). *Neuropsychological assessment* (4ª ed.). New York: Oxford University Press.
- Lima, C. F., Meireles, L. P., Fonseca, R., Castro, S. L. e Garrett, C. (2008). The Frontal Assessment Battery (FAB) in Parkinson's disease and correlations with formal measures of executive functioning. *Journal of Neurology*, 255(11), 1756-1761. doi: 10.1007/s00415-008-0024-6
- Matsuoka, T., Narumoto, J., Shibata, K., Okamura, A., Nakamura, K., Nakamae, ... Fukui, K. (2011). Neural correlates of performance on the different scoring systems of the Clock Drawing Test. *Neuroscience Letters*, 487(3), 421-425. doi: 10.1016/j.neulet.2010.10.069
- Ministério das cidades, Ordenamento do Território e Ambiente (2002). Decreto-Lei n.º 244/2002, de 5 de novembro. *Diário da República* — I Série A (255), 7101-7103.
- Murphy, K. R. e Davidshofer, C. O. (2004). *Psychological testing: Principles and applications* (6ª ed.). Englewood Cliffs, N.J.: Prentice-Hall.
- Nagahama, Y., Okina, T., Suzuki, N., Nabatame, H. e Matsuda, M. (2005). Neural correlates of impaired performance on the Clock Drawing Test in Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, 19(5-6), 390-396. doi: 10.1159/000084710
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113. doi: 10.1016/0028-3932(71)90067-4
- Paganini-Hill, A. e Clark, L. J. (2011). Longitudinal assessment of cognitive function by clock drawing in older adults. *Dementia and Geriatric Cognitive Disorders Extra*, 1(1), 75-83. doi: 10.1159/000326781
- Pallant, J. (2011). *SPSS survival manual: a step by step guide to data analysis using SPSS for Windows* (4ª ed.). Crows Nest NSW: Allen and Unwin.
- Patocskai, A. T., Pákási, M., Vincze, G., Fullajtár, M., Szimjanovszki, I., Boda, K., ... Kálmán, J. (2011). Is there any difference between the findings of clock drawing tests if

- the clocks show different times? *International Journal of Geriatric Psychiatry*, 26(4), 437-438. doi: 10.1002/gps.2501
- Peat, J (2002). *Health science research: A handbook of quantitative methods*. Crows Nest NSW: Allen and Unwin.
- Pinto, E. e Peters, R. (2009). Literature review of the Clock Drawing Test as a tool for cognitive screening. *Dementia and Geriatric Cognitive Disorders*, 27(3), 201-213. doi: 10.1159/000203344
- Raven, J., Raven, J. C. e Court, J. H. (2000). *Manual for Raven's Standard Progressive Matrices and Vocabulary Scales. Section 3: Progressive Standard Matrices*. San Antonio, TX: Harcourt Assessment.
- Razali, N. M. e Wah, Y. B. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of Statistical Modeling and Analytics*, 2(1), 21-33.
- Rey, A. (2002). Rey: *Teste de cópia de figuras complexas* (2ª ed.). Lisboa: CEGOC-TEA.
- Reznek, L. (2005). The Rey 15-item memory test for malingering: A meta-analysis. *Brain Injury*, 19(7), 539-543. doi: 10.1080/02699050400005242
- Seidl, U., Traeger, T. V., Hirjak, D., Remmele, B., Wolf, R. C., Kaiser, E., ... Thomann, P. A. (2012). Subcortical morphological correlates of impaired clock drawing performance. *Neuroscience Letters*, 512(1), 28-32. doi: 10.1016/j.neulet.2012.01.054
- Serra, A. V., Ponciano, E. e Relvas, J. (1982). Aferição da escala de autoavaliação de ansiedade, de Zung, numa amostra de população portuguesa - II. - Sua avaliação como instrumento de medida. *Psiquiatria Clínica*, 3(4), 203-213.
- Shua-Haim, J., Koppuzha, G., Shua-Haim, V. e Gross, J. (1997). A simple score system for clock drawing in patients with Alzheimer's disease. *American Journal of Alzheimer's Disease and Other Dementias*, 12(5), 212-215. doi: 10.1177/153331759701200505
- Shulman, K. I. (2000). Clock-drawing: is it the ideal cognitive screening test? *International Journal of Geriatric Psychiatry*, 15(6), 548-561. doi: 10.1002/1099-1166(200006)15:6<548::aid-gps242>3.0.co;2-u
- Shulman, K. I., Pushkar Gold, D., Cohen, C. A. e Zuccherro, C. A. (1993). Clock-drawing and dementia in the community: A longitudinal study. *International Journal of Geriatric Psychiatry*, 8(6), 487-496. doi: 10.1002/gps.930080606
- Shulman, K. I., Shedletsky, R. e Silver, I. L. (1986). The challenge of time: Clock-drawing and cognitive function in the elderly. *International Journal of Geriatric Psychiatry*, 1(2), 135-140. doi: 10.1002/gps.930010209

- Shoyama, M., Nishioka, T., Okumura, M., Kose, A., Tsuji, T., Ukai, S. e Shinosaki, K. (2011). Brain activity during the Clock-Drawing Test: Multichannel near-infrared spectroscopy study. *Applied Neuropsychology*, 18(4), 243-251. doi: 10.1080/09084282.2011.595450
- Simões, M., Sousa, L., Duarte, P., Firmino, H., Pinho, M. S., Gaspar, N., ... França, S. (2010). Avaliação da simulação ou esforço insuficiente com o Rey 15-Item Memory Test (15-IMT): Estudos de validação em grupos de adultos idosos. *Análise Psicológica*, 28(1), 209–226. doi: 10.14417/ap.269
- Slick, D. J., Tan, J. E., Strauss, E. H. e Hulstsch, D. F. (2004). Detecting malingering: a survey of experts' practices. *Archives of Clinical Neuropsychology*, 19(4), 465–473. doi: 10.1016/j.acn.2003.04.001
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643–662. doi: 10.1037/h0054651
- Takahashi, M., Sato, A., Nakajima, K., Inoue, A., Oishi, S., Ishii, T. e Miyaoka, H. (2008). Poor performance in Clock-Drawing Test associated with visual memory deficit and reduced bilateral hippocampal and left temporoparietal regional blood flows in Alzheimer's disease patients. *Psychiatry and Clinical Neurosciences*, 62(2), 167-173. doi: 10.1111/j.1440-1819.2008.01750.x
- Tranel, D., Rudrauf, D., Vianna, E. P. M. e Damasio, H. (2008). Does the Clock Drawing Test have focal neuroanatomical correlates? *Neuropsychology*, 22(5), 553-562. doi: 10.1037/0894-4105.22.5.553.
- Tuokko, H., Hadjistavropoulos, T., Rae, S. e O'Rourke, N. (2000). A comparison of alternative approaches to the scoring of Clock Drawing. *Archives of Clinical Neuropsychology*, 15(2), 137–148. doi: 10.1016/S0887-6177(99)00003-7
- Tuokko, H. e O'Connel, M. E. (2006). A review of quantified approaches to the qualitative assessment of clock drawing. Em A. M. Poreh (Ed.) *The Quantified process approach to Neuropsychological Assessment* (pp. 173-206). New York: Taylor and Francis.
- Writer, B. W., Schillerstrom, J. E., Regwan, H. K. e Harlan, B. S. (2010). Executive clock-drawing correlates with performance-based functional status in people with combat-related mild traumatic brain injury and comorbid posttraumatic stress disorder. *Journal of Rehabilitation Research and Development*, 47(9), 841-850. doi: 10.1682/jrrd.2009.10.0162